

# Assessing and Refining Group Take-Home Exams as Authentic, Effective Learning Experiences

By Corey M. Johnson, Kimberly A. Green, Betty J. Galbraith, and Carol M. Anelli

*The learning goals of a lower division honors course, Science as a Way of Knowing, include critical thinking, scientific literacy, quantitative reasoning, communication, and teamwork. To help students develop skills and competencies for the course learning outcomes, we used a case study and developed scaffolded activities and assignments that targeted discipline-relevant tasks, for example, primary literature search, evaluation of source credibility, hypothesis construction, data interpretation, and restatement of scientific content into lay terminology. We then implemented group take-home exams, which feature rigorous, open-ended questions in authentic contexts, requiring students to apply knowledge and competencies cooperatively to new situations. Data from five semesters show that, in comparison to traditional exams, many students feel that group take-home exams reduce test anxiety, foster interpersonal skills, are more rigorous, and better enable them to apply and synthesize knowledge and deepen their comprehension of the subject matter. Our study augments research on group exams that use an open-ended response format.*

Advocates for improving undergraduate science education call for pedagogies that engage students in relevant, “real-world” problem solving and cooperative learning (American Association for the Advancement of Science, 2011; Handelsman et al., 2004). From his literature review, including a meta-analysis of 164 studies of cooperative learning methods, Michael (2006, p. 162) concluded that “little doubt [exists] that students in groups learn more.” Studies also report the effectiveness of group exams as a learning tool. Group exams that require critical and higher order thinking skills deepen contextualization and improve retention (Drouin, 2010; Michael, 2006; Zipp, 2007). Challenging exam questions promote student discussion, fostering communication skills (oral, written, listening) and facilitating learning, in part through students serving as peer instructors (Michael, 2006; Simkin, 2005). Group exams can reduce test anxiety (Morgan, 2005; Simkin, 2005; Zipp, 2007) and enable students to practice interpersonal skills such as collaboration, reciprocity, team building, troubleshooting, leadership, conflict resolution, and trust (Rao, Collins, & DiCarlo, 2002; Simkin, 2005; Zipp, 2007). Students can build on one another’s academic and personal strengths and demonstrate enhanced motivation

and engagement, behaviors driven by feelings of responsibility to the group (Cortright, 2003; Zipp, 2007).

Educators and learning experts increasingly view fixed-choice exams as limited measures of student learning as they create artificial situations that do not reflect learners’ responses in real-world situations (Oakleaf, 2008; Simkin, 2005). Similarly, hourly or “midterm” exams fall short because they impose unrealistic time limits and often do not target higher level cognitive skills. In contrast, performance-based tasks that simulate real-life application of skills, knowledge, and competencies enable assessment in authentic contexts (Mueller, 2012; Oakleaf, 2008). By emphasizing what students can *do* with what they know, authentic assessment complements traditional approaches that emphasize what students *know* about a body of knowledge.

The literature on group work, together with that on authentic assessment, suggests that group take-home exams could provide both a valuable learning experience and a means to assess student performance on real-world tasks. The authors had the opportunity to experiment with such exams when our honors college adopted a new curriculum in 2008, in which introductory courses focus on scholarly literature use in preparation for the required thesis. (Honors courses fulfill general education

requirements for the baccalaureate degree.) We were asked to develop a new course, *Science as a Way of Knowing*, emphasizing information and scientific literacies, critical thinking, quantitative reasoning, and communication. We hypothesized that group take-home exams, which relieve the time constraint of hourly exams, could provide real-world problems while fostering learning and cooperative work. Our literature review yielded no published reports on exams like ours, which student groups complete entirely outside of class and which feature authentic, discipline-relevant tasks.

Here we report our findings on student attitudes and perceptions and our research on implementation and refinement of group take-home exams used in lieu of midterm or hourly exams. Our exams feature rigorous, open-ended questions in authentic contexts and require students to apply knowledge and competencies cooperatively to new situations. Such exams embody four of seven long-recognized essentials of good practice in undergraduate education: time on task, communication of high expectations, active learning techniques, and reciprocity and cooperation (Chickering & Gamson, 1987). In addition, good practice for effective assessment involves representatives from across the educational community (American Association for Higher Education and Accreditation, 1996). Our team comprises a subject expert (professor and instructor of record), an assessment specialist, and two instruction librarians, one specializing in science.

### Methods of instruction and assessment

Nonscience majors constituted 20%–25% of course enrollment,

with class size limited to 25 lower division students. We used backward design (Wiggins & McTighe, 2005) to align course learning goals and outcomes (contextualized within evolutionary theory and the history of biology; sample syllabus, Appendix A, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)), created appropriate activities, and provided timely written feedback (Fink, 2003). Earlier assessment efforts (Johnson, Anelli, Galbraith, & Green, 2011) guided our pedagogy. Interactive lectures and discussions focused on scientific database use, types of sources, credibility of sources and expertise, and interpretation and evaluation of research articles. To move them progressively through skills and competencies emphasized on exams, we had students work in small groups (three to five members each), in and outside of class, to complete scaffolded assignments that included a case study (Hollister, 2005) and several question sets (samples, Appendix B, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)), most of which scrutinized Moran (2004), a primary article selected for its rigor and design (i.e., test of predictions of evolutionary theory using several experimental approaches), minimal use of disciplinary jargon, and conceptual accessibility. Gilen (2007) served as a supplemental “how to” guide on the structure, stylistic conventions, and critical reading of primary scientific articles.

For group take-home exams, the instructor selected research articles to which students had not been previously exposed and whose subject matter had not been discussed. The exams featured a mix of brief essays and short answers, requiring students to search the primary literature (con-

temporary and historical), interpret data, and restate scientific content and theories into lay terminology (sample exam, Appendix C, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)).

Because our students comprise a spectrum of backgrounds, majors, and interests, for each exam the instructor assembled students into small groups, comparably mixed (gender, students’ self-reported coursework, quality of student’s individual work to date, major and experience; Siciliano 2001). The instructor assembled new groups for each exam. Within the course online learning management system (LMS), each group had its own workspace, accessible only to its student members and the current authors. On the day that students gained access to the exam, the instructor allowed 10–15 minutes of class time for groups to meet and make organizational plans. The instructor exhorted students to be “good team members” and to contact her with unresolvable group issues/questions. She monitored groups’ progress informally by querying groups during class or via LMS; clarification issues were resolved at once. Most groups did their own troubleshooting, but the instructor intervened at her discretion when asked. With the exam, the instructor posted a dissension form for students who wished to submit their own answer(s) if they disagreed with their groups’ answer(s); in our five semesters of teaching the course, no student ever used the form.

When the deadline for the completed exam expired, the instructor posted the answer key (sample grading key, Appendix D, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)) to the LMS site. She devoted the next class period to discussion of the

exam, facilitated by visual projection of the grading key, and emphasized problematic questions as revealed by student feedback. Discussion sometimes led to improvements in how future exams were written (e.g., providing detailed breakdown of point allocations for exam questions; also see Appendix E, [www.nsta.org/](http://www.nsta.org/)

[college/connections.aspx](http://college/connections.aspx)).

We typically administered two group take-home exams per semester, in Weeks 6 and 11 of a 15-week semester. Groups had 7–10 days for each exam, each of which contributed  $14\% \pm 2\%$  ( $M \pm SD$ ) to the overall course grade, depending on semester. To encourage groups

to take ownership of their working relationship, individual effort was not factored into group exam grades; within a given group, all members received the same exam grade. All students took an individual, in-class, closed-book final exam, worth  $16\% \pm 3\%$  ( $M \pm SD$ ) of the overall course grade, depending on semester. Individual effort (attendance, various assignments; see syllabus, Appendix A, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)) contributed  $48\% \pm 4\%$  ( $M \pm SD$ ) to a student's overall course grade.

To assess students' attitudes, background knowledge, and performance, we used pre- and postcourse anonymous questionnaires (Appendix F, available at [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)), honors college anonymous end-of-course evaluations, and exam scores (generated by the instructor using grading keys; see sample, Appendix D, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)). We did not track individual students in pre–post pairings. For Likert-scale pre- and postcourse questionnaire prompts, which did not change from semester to semester, we pooled data from all five semesters that we taught the course: fall 2008 (initial course offering), spring 2009, fall 2010, fall 2011, and fall 2012. For open-ended prompts as to the positive and negative aspects of group exams, we coded and summarized data by semester and across all five semesters. Beginning with fall 2010, students also completed self- and peer-performance forms after each exam (based on criteria in Isaacs, 2002), which the instructor used to assess group dynamics (student grades were not impacted). Our institutional review board approved our protocols and instruments prior to implementation.

**FIGURE 1**

**Template for small group take-home exam contract**

**GROUP CONTRACT FOR EXAM # \_\_\_\_\_**  
(COURSE NAME & NUMBER, DATE)

**GROUP NAME \_\_\_\_\_**

We, the undersigned, have together devised and agreed to an initial plan (below) for working on the exam (*insert initial deadline(s) for work to be shared via Google docs*). We will indicate individual contributions (tracking, color coding, etc.) and contact (*instructor name*) ASAP with concerns.

**Plan:** (*Group inserts details here*)

Each of us also agrees to do the following:

1. Attend group meetings (virtually or in person)
2. Maintain contact w/group members
3. Communicate constructively to group discussion & answers
4. Be cooperative and understanding
5. Take a leadership role as needed
6. Encourage and assist my team members
7. Complete all tasks agreed upon by the group *on time*
8. Complete/upload my exam portion and share it with the team by (*date, time*)
9. Read, comment on, and edit the **entire exam** by the time agreed upon by the group
10. Ensure that the exam **final version** is ready to be uploaded by (*date, time*)
11. Notice and work to curtail whatever tendencies I may sometimes exhibit that others may perceive as uncooperative

If any one of us causes difficulty with the group, and/or breaks the contract in any way, we understand that the other team members have the right and are expected to contact (*instructor name*) and inform her of the situation. We further understand that (*instructor name*) will serve as arbiter and may decide to penalize the teammate in question by lowering his/her grade in accord with the situation, or making the teammate complete the exam alone.

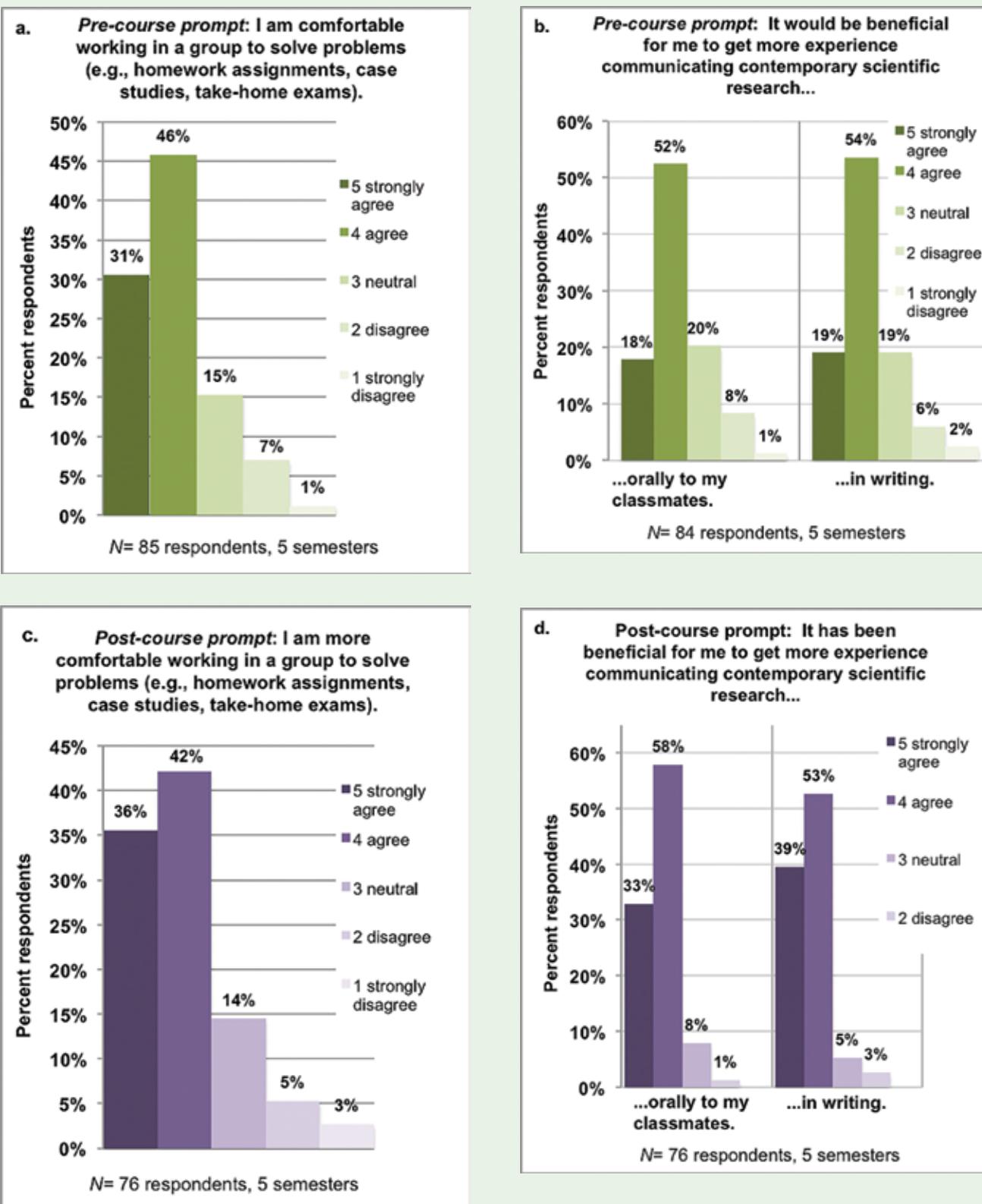
Signed,

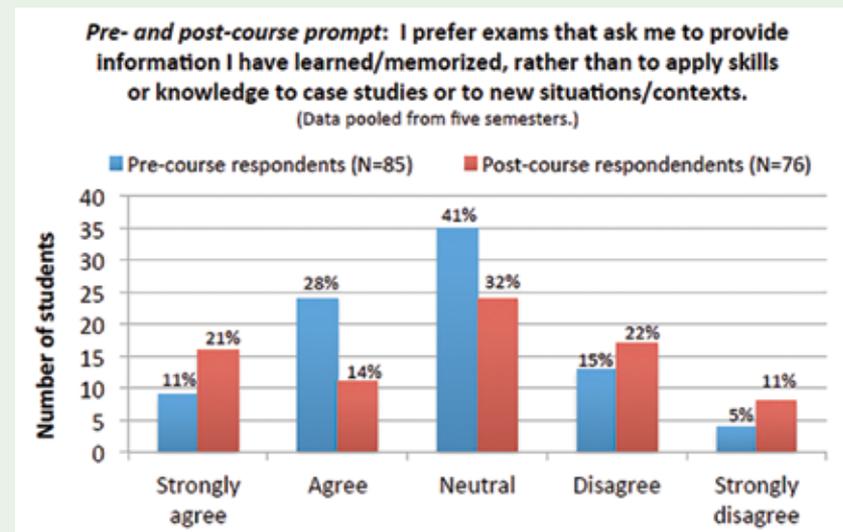
\_\_\_\_\_

\_\_\_\_\_

**FIGURE 2**

Pre- and postcourse questionnaire responses regarding group work and communication (Likert scale responses).



**FIGURE 3****Comparison of exam-type preference, pre- vs. postcourse questionnaires.**

Our experiences and formative assessment over five semesters led us to implement changes to improve the group-exam process: (a) designate a group leader to keep members on task and assemble their work (we assigned leaders randomly; no student served more than once; implemented with Exam 1, fall 2011); (b) schedule the exam over a 10-day period (instead of 7), spanning two weekends (ours is a rural campus and students' co-curricular activities often occur at distant locations; this change mitigated stress; implemented with Exam 1, fall 2011); (c) require groups to submit a completed contract detailing plans for exam completion and participation, with frequent updates (implemented with Exam 1, fall 2012; Figure 1); and (d) implement a draft deadline on Day 7 for groups' answers to be uploaded to their LMS site (implemented with Exam 2, fall 2012). Although these changes were important to the group exam experience, the fundamental

structure of the process remained the same throughout the five semesters. Assessment of improvement based on these specific changes is not possible because, given our enrollment, we taught only one course section per semester and had no "nontreatment" group.

## Results

### *Pre- and postcourse anonymous questionnaires*

At the course outset, the majority of students indicated that they were comfortable working in groups, and most appeared motivated to gain experience in oral and written communication of scientific research (Figure 2a and b). Most postcourse respondents felt they had made gains in these areas (Figure 2c and d).

We asked students if they preferred exams that ask for information learned/memorized or exams that require the application of skills or knowledge. The data did not indicate a statistically

significant shift in the means, but there was a significant difference in the distribution of responses between pre- and postcourse observations (cross-tabulation analysis:  $\chi^2(4) = 10.24, p < .05$ ; Figure 3). At the course outset, 41% of students expressed neutrality on exam preference versus 32% in postcourse analyses. Apparently a comparable percentage of students pre- versus postcourse preferred learned/memorized exams (39% vs. 35%), whereas the percentage of students who dissented from that preference increased by 13% in pre- versus postcourse data (33% vs. 20%).

A set of six questions in the post-course questionnaire targeted students' perceptions of our group take-home exams compared with traditional exams (hourly, in-class, completed individually). Fifty-five percent of respondents indicated that group take-home exams were more rigorous than traditional exams, compared with 15% who dissented from that view (31% being neutral); student perceptions were split on whether take-home exams required more (40%) or less (35%) individual effort than traditional exams (Figure 4a). The vast majority reported that compared with hourly exams, group take-home exams enabled them to apply and synthesize knowledge, deepened their comprehension, benefitted their interpersonal skills, and increased their awareness of course relevance (Figure 4b).

### Course activities and resources

Our postcourse questionnaire included a list of course activities and resources from which we asked students to select those that helped them learn. Of 77 respondents, 71% ( $N = 55$ ) selected the entry "group work for exams," making it the fourth

most selected item after PowerPoints (84%), class discussions (84%), and the textbook (80%; 5 semesters; data not shown).

### Comments on group take-home exams

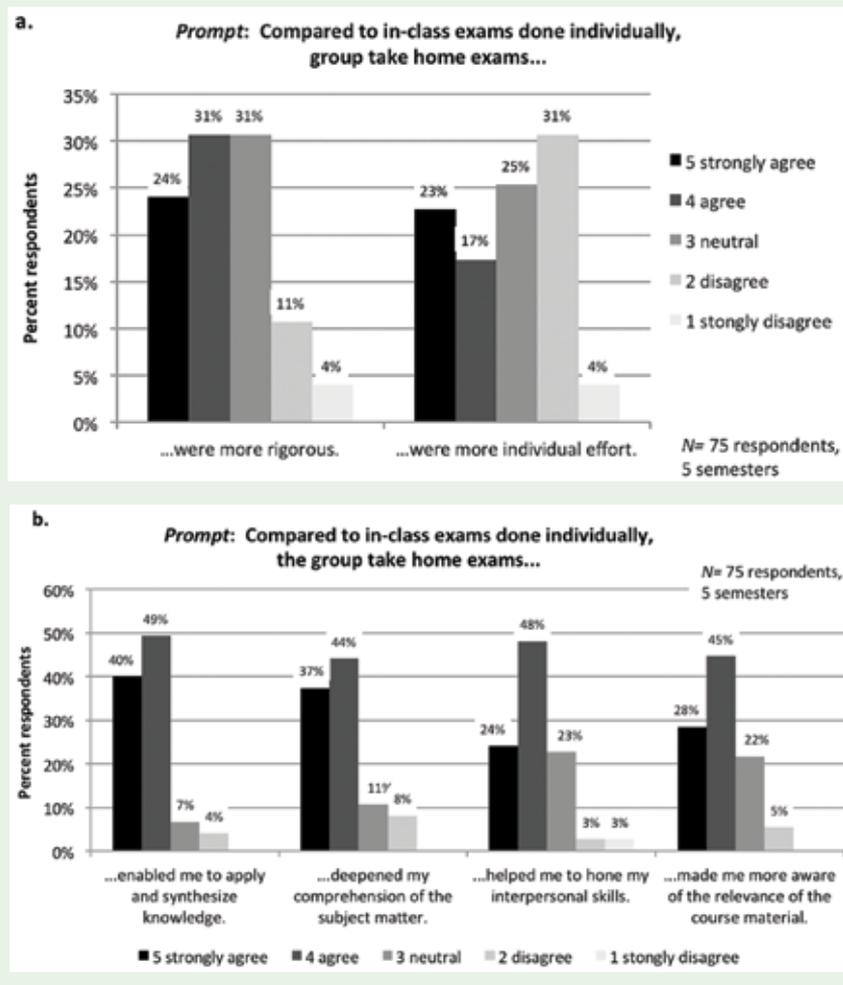
Postcourse questionnaires prompted students to “comment on the positive and/or negative aspects of group take-home exams.” In five semesters’ data, 43 of 77 respondents (56%) yielded a pool of 32 negative and 41 positive comments, which we categorized and summarized (Table 1). Most negative comments focused on group members’ unequal contributions to the exam (quality and/or quantity), followed by exam length/workload. The largest number of positive comments cited various benefits of the group experience, followed by learning quality afforded by group exams. Analysis showed no significant trend in positive-to-negative comment ratios in successive semesters (logistic regression:  $\chi^2(1) = 1.73, p = .19$ ). The largest positive-to-negative comment ratios occurred in the last two semesters, by which time some or all improvements to the exam process (see Methods section) had been implemented.

### Honors course evaluations

Students completed honors college anonymous course evaluations for three semesters (fall 2010 to fall 2012). Respondents (58.6% response rate,  $N = 58$ ) indicated that they had improved at collaborating with classmates (62.1% *strongly agree* / 32.8% *agree*) and that the course had positively impacted their skills in critical thinking (64% *strongly agreed* / 36% *agreed*), writing (33% *strongly agreed* / 59% *agreed*), and quantitative reasoning (53% *strongly agreed*

**FIGURE 4**

**Postcourse questionnaire responses regarding group take-home exams.**



/ 36% *agreed*; data not shown).

### Exam scores

Group take-home exams required transfer of skills in information and science literacies, critical thinking, quantitative reasoning, and communication about science to new situations (sample exam, Appendix C, [www.nsta.org/college/connections.aspx](http://www.nsta.org/college/connections.aspx)). On our grading scale, group exam grades averaged B<sup>+</sup> to A<sup>-</sup> ( $M = 88.3\% \pm 3.7\% SD$  for 10 exam averages; data

not shown). For informational (not comparative) purposes, final exam grades (closed book, completed by individual students) over five semesters averaged B to B<sup>+</sup> ( $M = 82.3\% \pm 3.0\% SD$ ; range of averages for 5 semesters = 76.8% – 85.6%). No exams were graded on a “curve.”

### Postexam self- and peer-performance forms

Students rated their own and each group member’s performance on ex-

ams according to 12 criteria using a Likert scale (1 = *very well*, 2 = *adequate*, 3 = *poor*). Eighty-six students completed self-performance forms, yielding a total of 1,032 ratings (86\*12; data not shown). The vast majority of students rated their own efforts highly (*very well* = 815 [79%], *adequate* = 183 [18%], *poor* = 23 [2%], *no response* = 11 [1%]). The criteria with the most 3s (*poor*) were “I took a leadership role as needed” and “I read and commented in a timely manner on drafts of the exam.” Peer-performance forms ( $N = 260$ ) yielded 3,120 ratings (260\*12), with aggregate ratings mirroring those for self-performance forms (*very well* = 2,609 [84%], *adequate* = 415 [13%], *poor* = 87 [3%],

*no response* = 9 [ $<1\%$ ]). The criteria with the most 3s (*poor*) were “Read and commented in a timely manner on drafts of the exam” and “Encouraged and assisted other group members.”

### Comments on group take-home exams

Performance forms prompted students to state positive and negative things about the group exam and suggest improvements. Of 81 students who completed the forms, 68 (83%) responded to the prompts, yielding a total of 41 negative and 85 positive comments (Table 2). To categorize and summarize students’ comments, we used the same criteria as those for Table 1. In contrast

to postcourse questionnaire data in Table 1, the largest number of negative student comments postexam focused on group dynamics (e.g., problems with scheduling, communication, leadership, peers’ work habits/behavior), followed by general/specific criticism of the exam (e.g., word limits, certain exam questions, challenges dividing workload). The largest number of positive comments cited greater time to complete the group take-home versus an in-class exam, followed by benefits of the group experience (e.g., collaboration/brainstorming, greater exposure to learning styles, practice with conflict resolution, authenticity of tasks, sharing of workload, knowing classmates bet-

**TABLE 1**

**Student response rates on postcourse questionnaires, with distribution and coding of open-ended comments about the group take-home exams.**

Semester	Total students enrolled	Students who completed questionnaire	Students who responded to prompt	NEGATIVE COMMENTS (codings below)						POSITIVE COMMENTS (codings below)						
	N	N (%)	N (% of respondents)	#1	#2	#3	#4	#5	SUM	#1	#2	#3	#4	#5	#6	SUM
Fall '08	16	11 (69%)	8 (73%)	2	3	1	1	1	8	6	0	1	1	0	0	8
Spring '09	17	7 (41%)	3 (43%)	2	0	0	0	1	3	0	1	1	0	0	0	2
Fall '10	20	13 (65%)	11 (85%)	2	2	3	1	0	8	0	4	1	2	0	0	7
Fall '11	22	22 (100%)	13 (59%)	3	2	1	2	0	8	5	2	4	2	0	0	13
Fall '12	24	24 (100%)	8 (33%)	3	1	0	1	0	5	3	3	1	1	2	1	11
<b>Total</b>	<b>99</b>	<b>77 (75% ± 25%)</b>	<b>43 (56% ± 21%)</b>	<b>12</b>	<b>8</b>	<b>5</b>	<b>5</b>	<b>2</b>	<b>32</b>	<b>14</b>	<b>10</b>	<b>8</b>	<b>6</b>	<b>2</b>	<b>1</b>	<b>41</b>
			<b>Percent</b>	<b>38</b>	<b>25</b>	<b>16</b>	<b>16</b>	<b>6</b>	<b>100%</b>	<b>34</b>	<b>24</b>	<b>20</b>	<b>15</b>	<b>7</b>	<b>2</b>	<b>100%</b>

#### # Negative comment codings

1. Unequal group sizes and individual contributions
2. Worse than in-class exams
3. Doesn't prepare for in-class, individual final exam
4. Group dynamics issues
5. Criticism of exam, general and specific

#### # Positive comment codings

1. Benefits of group collaboration
2. Conducive to quality learning
3. Praise for exam, general and specific
4. Better than in-class exams
5. More time to complete vs. in-class exam
6. Requirement for draft exam on day 7

Note: An individual student may have offered >1 comment, both positive and negative.

ter, working harder for the group). The ratio of positive-to-negative comments trended toward increasing over the last three semesters, although this trend was not significant (logistic regression analysis:  $\chi^2(1) = 1.63, p = .20$ ); most positive comments for fall 2012 Exam #2 (32%,  $N = 10$ ) cited implementation of the Day 7 draft deadline.

## Discussion

The abilities to evaluate and make sense of scientific findings and to interact productively with others represent critical life skills and embody our course learning outcomes. We were eager to explore group take-home exams as impactful student learning experiences and as tools

to gauge students' performance, but cognizant of challenges. Honors students are academically motivated and have good track records as independent achievers, as reflected in their scores on the cumulative final exam (which included higher order questions), and many decry group work. In addition, nonmajors' attitudes toward science can dampen their interest, and some students in our lower division course are still adjusting to the demands of college coursework.

Challenges notwithstanding, our results suggest that group take-home exams can provide positive learning experiences and simultaneously facilitate assessment and evaluation of students' ability to apply science

knowledge and skills collaboratively to authentic, real-world tasks. Initially our students expressed anxiety that their semester exam grades would reflect their group's collective efforts, partly because they lacked experience with this type of exam, but also because, as highly engaged individuals, their busy lives presented scheduling difficulties. Yet most students left the course feeling more comfortable about performing group work, and the majority felt that group take-home exams enabled them to apply and synthesize knowledge, deepened their comprehension of the course material, made them more aware of its relevance, and helped them hone interpersonal skills. The vast majority also indicated that our course posi-

**TABLE 2**

**Student response rates on postexam performance forms, with distribution and coding of open-ended comments about the group exam.**

Semester	Exam	Total students enrolled N	Students who completed evaluation form N (%)	Students who responded to prompt N (% of respondents)	NEGATIVE COMMENTS (codings below)						POSITIVE COMMENTS (codings below)						
					#1	#2	#3	#4	#5	SUM	#1	#2	#3	#4	#5	#6	SUM
Fall '10	#2	20	20 (100%)	16 (80%)	0	4	1	4	2	11	3	4	1	1	7	0	16
Fall '11	#1	22	19 (86%)	11 (58%)	0	2	0	6	1	9	10	2	4	1	2	0	19
Fall '12	#1	24	21 (87%)	20 (95%)	0	0	0	6	5	11	2	3	1	3	10	0	19
Fall '12	#2	24	21 (87%)	21 (100%)	3	1	0	2	4	10	5	2	5	3	6	10	31
<b>Total</b>	<b>4</b>	<b>90</b>	<b>81 (90% ± 7%)</b>	<b>68 (83% ± 19%)</b>	<b>3</b>	<b>7</b>	<b>1</b>	<b>18</b>	<b>12</b>	<b>41</b>	<b>20</b>	<b>11</b>	<b>11</b>	<b>8</b>	<b>25</b>	<b>10</b>	<b>85</b>
<b>Percent</b>					<b>7</b>	<b>17</b>	<b>3</b>	<b>44</b>	<b>29</b>	<b>100%</b>	<b>24</b>	<b>13</b>	<b>13</b>	<b>9</b>	<b>29</b>	<b>12</b>	<b>100%</b>

### # Negative comment codings

1. Unequal group sizes and individual contributions
2. Worse than in-class exams
3. Doesn't prepare for in-class, individual final exam
4. Group dynamics issues
5. Criticism of exam, general and specific

### # Positive comment codings

1. Benefits of group collaboration
2. Conducive to quality learning
3. Praise for exam, general and specific
4. Better than in-class exams
5. More time to complete vs. in-class exam
6. Requirement for draft exam on day 7

Note: An individual student may have offered >1 comment, both positive and negative.

tively affected their critical thinking, writing, quantitative reasoning, and science communication skills (Anelli, Johnson, Galbraith, & Green, 2015). As group take-home exams and related group work constituted a considerable portion of course activities and final grade computation, these findings lend support to the positive impact of cooperative work. Our exams did not compromise rigor; groups generally performed well but seldom achieved 100%, and about half of our students felt our exams were more rigorous than traditional exams.

Students' open-ended comments provide insight into their perceptions of group take-home exams. Some students expressed excitement that the exams required critical thinking and application of knowledge, and some characterized the exams as a great way to learn and retain information, practice concise writing, have fruitful discussions, and learn conflict resolution. Others saw the group take-home format as superior because it relieved the pressure from individual performance and "cramming" for exam preparation and/or completion. One of the students described our exams as "fun," another claimed to "love them," yet another called it "one of [his/her] best experiences for college exams."

As anticipated, based on the literature and our teaching experiences, challenges surfaced and students voiced complaints and concerns. We strove to address issues proactively, and as students will undoubtedly encounter undesirable peer behaviors throughout their lives, we do not see challenges as reason to avoid group exams. Interestingly, although the types of positive and negative comments made on questionnaires versus performance forms were

comparable, the relative rankings of comments differed between the two types of data. The top two negative comments on postcourse questionnaires focused on unequal group sizes/individual contribution issues and the group exam format, compared with group dynamics issues and criticism of the exam itself on postexam performance forms. Similarly, the top positive comment category on postcourse questionnaires was "benefits of group collaboration," whereas on postexam forms, "greater time allowance" was the most popular commendation, accounting for 30% of student comments compared with 5% on postcourse questionnaires. Not to overemphasize these data, but they do suggest that instructors should consider that students' feedback can vary over time. Perhaps when the exam experience is relatively fresh, students value most the greater time allowance for take-home exams, but with the passage of time they may recall both the benefits of teamwork and their discontent with particular teammates. Teaching and assessment expert Maryellen Weimer (2002) urged instructors not to ask students whether they "liked" a particular activity:

That is an irrelevant criterion . . . The questions you need answered are these: "How did that activity. . . affect your learning?" "What about it needs to change so that if we do it again, you will learn more?" (p. 199)

Almost three quarters of our students selected "work for group exams" as a course activity that helped them learn, underscoring group take-

home exams as impactful learning experiences. In the last two semesters that we taught the course, on both postexam performance forms and postcourse questionnaires, we received more positive than negative comments, suggesting that the assessment-based changes we implemented helped improve the exam experience.

It does appear that group take-home exams were on students' minds, as it was the most prolific topic for comment among postcourse respondents who provided "one or two concrete suggestions" to improve the course. Many students wanted at least one take-home exam to be a solo effort; this would have significantly increased our grading effort and diminished the opportunity for students to hone collaborative skills. Others suggested that we account for individual effort by giving students individual grades plus the group's grade, an approach many instructors use (Simkin, 2005; Rao et al., 2002); we chose not to do so because we wanted to incentivize students to improve at working cooperatively. Still other students believed that the group exam ill-prepared them for the final exam, administered all five semesters as an individual, closed-book, in-class exam. To address this concern we provided a study sheet that showed point allocations for the various topics and types of questions that would appear on the final. On the basis of five semesters' worth of data, students earned on average 5 percentage points less on the final exam versus group take-home exams. We embedded questions in the final that required application of knowledge and skills per the syllabus, and students' performance and their self-reported gains suggest

that course learning outcomes were adequately met (Anelli et al., 2015).

Most groups functioned well, even excellently (e.g., one student wrote, “I love my group!”), but we instituted changes to diminish procrastination and project management issues. Our Day 7 draft policy with a Day 10 exam deadline garnered many positive responses and improved collaboration. Students used the “extra” time to review and edit teammates’ answers and complained less about the perceived need to “divide-and-conquer” exam questions, which they felt negatively impacted their learning. We concur, and through assessment we discovered that our students needed practice developing the habit of reviewing teammates’ answers. Assessment also led us to designate a leader for each group, causing “lack of leadership” complaints to disappear. Technological advances also facilitated group work. In 2008, students had difficulties communicating with group members and tracking the latest version of their group’s exam answers; both problems diminished with the greater prevalence of mobile phones, texting, and students’ familiarity with and use of Google Docs.

For instructors, group take-home exams offer pluses and minuses. The format affords development of complex, in-depth questions that can draw on an array of materials: the current scientific literature, online archival materials, the *New York Times* Tuesday science section, scholarly essays, YouTube videos, etc. Taking advantage of this flexibility, we designed creative (dare we say “fun”?) exams that were amenable to assessment of student achievement of outcomes. Plagiarism concerns were essentially nonexistent because answers could not be “googled,” and

many pairs of student eyes reviewed each exam answer for proper idea attribution. We monitored student progress, conceptual blocks, and/or misperceptions by informally asking how the exam was going. We could assess each student’s contribution by having groups indicate in the e-version of their exam who contributed what. Having only 5 or 6 exams to grade versus 25 represented a “plus.” Preparing a grading key with detailed point allocations before evaluating groups’ answers promoted instructor objectivity, saved grading time, and helped provide feedback during discussion of the exam. For large enrollment courses with significant grading demands, O’Dowd (2011) provided guidance on implementing a choice of fixed answers for questions that target higher levels of understanding.

On the negative side, we found that developing a group take-home exam required more effort than a traditional hourly exam as did writing a student-friendly grading key for open-ended responses. Locating research articles that were relatively free of scientific jargon yet amenable to testing students’ skills (as opposed to specific content knowledge), and creating exam questions aimed at the application of scientific competencies, took time. Scientific research in the media (including the “Ig Nobel Prizes”) often provided leads to provocative, suitable articles.

### Conclusions and recommendations

On balance, our experiences with group take-home exams lead us to conclude that the positive aspects outweigh the negative. Students can practice cooperative skills and perform discipline-relevant tasks, and many come to view the exams as a

positive learning experience.

We have not found a published report on exams that features authentic, discipline-relevant tasks for student groups to complete entirely outside of class. We believe instructors can use such exams to measure and enrich student learning and transferability of science knowledge and competencies, a need increasingly drawing the attention of educators (Mervis, 2013). We offer these recommendations. First, build motivation: Articulate your course learning goals to your students and explain how their attainment of those goals will be demonstrated by their performance. Second, build skills and give practice: Have students practice teamwork and self/peer evaluations on scaffolded, low-stakes activities, for example, homework assignments worth minimal points, before administering a group take-home exam (Morgan, 2005). Third, assess and guide: Assess students’ collaborative and work skills and provide prompt guidance and intervention as needed; give feedback regularly. Fourth, be clear and fair: State your exam policies and implement them consistently and fairly. Finally, refine: Use assessment recurrently to make needed adjustments to your exams and policies. ■

### Acknowledgments

*We acknowledge Dr. Erica Suchman, Colorado State University, whose 2008 presentation on group exams for her microbiology course inspired the instructor to experiment with and expand on this pedagogy, and Dr. David Sloan Wilson, Binghamton University, who suggested the use of contracts for our exams.*

### References

American Association for the Advancement of Science. (2011).

- Vision and change in undergraduate biology education: A call to action.* Washington, DC: Author. Retrieved from <http://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf>
- American Association for Higher Education and Accreditation. (1996). *Nine principles of good practice for assessing student learning.* Washington, DC: Author. Retrieved from [http://assessment.uconn.edu/docs/resources/AAHE\\_Principles\\_of\\_Good\\_Practice.pdf](http://assessment.uconn.edu/docs/resources/AAHE_Principles_of_Good_Practice.pdf)
- Anelli, C. A., Johnson, C. M., Galbraith, B. J., & Green, K. A. (2015). *Using group take-home exams to develop collaborative learning, scientific literacy, and higher-order cognitive skills.* Unpublished manuscript.
- Chickering, A. M., & Gamson, Z. A. (1987). Seven principles for good practice in undergraduate education. *American Association for Higher Education Bulletin*, 120, 3–7.
- Cortright, R. N. (2003). Student retention of course content is improved by collaborative group testing. *Advances in Physiology Education*, 27, 102–108.
- Drouin, M. (2010). Group-based formative summative assessment relates to improved student performance and satisfaction. *Teaching of Psychology*, 37, 114–118.
- Fink, L. D. (2003). *A self-directed guide to designing courses for significant learning.* San Francisco, CA: Jossey-Bass.
- Gillen, C. M. (2007). *Reading primary literature: A practical guide to evaluating research articles in biology.* San Francisco, CA: Pearson/Benjamin Cummings.
- Handelsman, J. D., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., . . . Wood, W. B. (2004). Scientific teaching. *Science*, 304(5670), 521–522.
- Hollister, C. (2005). *Rising temperatures, differing viewpoints: A case study on the politics of information.* Buffalo, NY: University of New York at Buffalo's National Center for Case Study Teaching in Science. Retrieved from <http://sciencecases.lib.buffalo.edu/cs/index.asp>
- Isaacs, G. (2002). *Assessing group tasks.* Queensland, Australia: University of Queensland, Teaching and Educational Development Institute. Retrieved from <http://www.itl.usyd.edu.au/assessmentresources/pdf/Link11.pdf>
- Johnson, C. M., Anelli, C. A., Galbraith, B. J., & Green, K. A. (2011). Information literacy instruction and assessment in an honors college science fundamentals course. *College and Research Libraries*, 72, 533–547.
- Mervis, J. (2013). Transformation is possible if a university really cares. *Science*, 340(6130), 292–295.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4), 159–167.
- Moran, A. (2004). Egg size evolution in tropical American arcid bivalves: The comparative method and the fossil record. *Evolution*, 58, 2718–2733.
- Morgan, B. M. (2005). Cooperative learning in higher education: A comparison of undergraduate and graduate students' reflections on group exams for group grades. *Journal on Excellence in College Teaching*, 16, 79–95.
- Mueller, J. F. (2012). *What is authentic assessment?* Retrieved from <http://jfmueller.faculty.noctrl.edu/toolbox/whatisit.htm>
- Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy assessment approaches. *Portal: Libraries and the Academy*, 8, 233–253.
- O'Dowd, D. (2011). *How to write good multiple-choice exam questions.* Retrieved from <http://pdffooz.org/k-28158359.html>
- Rao, S. P., Collins, H. L., & DiCarlo, S. E. (2002). Collaborative testing enhances student learning. *Advances in Physiology Education*, 26, 37–41.
- Siciliano, J. (2001). How to incorporate cooperative learning principles in the classroom: It's more than just putting students in teams. *Journal of Management Education*, 25(8), 8–20.
- Simkin, M. G. (2005). An experimental study of the effectiveness of collaborative testing in an entry-level computer programming class. *Journal of Information Systems Education*, 16, 273–280.
- Weimer, M. (2002). *Learner-centered teaching: Five key changes to practice.* San Francisco, CA: Jossey-Bass.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Upper Saddle River, NJ: Pearson Education.
- Zipp, J. T. (2007). Learning by exams: The impact of two-stage cooperative tests. *Teaching Sociology*, 35, 62–76.

---

**Corey M. Johnson** ([coreyj@wsu.edu](mailto:coreyj@wsu.edu)) is head, Library Instruction Team; **Kimberly A. Green** is director, Office of Assessment and Teaching; and **Betty J. Galbraith** is science librarian and instruction coordinator, all at Washington State University in Pullman. **Carol M. Anelli** is a professor and associate chair in the Department of Entomology at The Ohio State University in Columbus.

---